FIR H^∞ EQUALIZATION

Alper T. Erdogan — ISL Labratory, EE Dept., Stanford University, CA 94305

Babak Hassibi — Bell Labs Innovations, Lucent Technologies, Murray Hill, NJ07974-0636

Thomas Kailath — ISL Labratory, EE Dept., Stanford University, CA 94305

ABSTRACT

We approach FIR equalization problem from an H^∞ perspective. First, we formulate the calculation of the optimal H^∞ performance for a given equalization setting as a semidefinite programming (SDP) problem. H^∞ criterion provides a set of FIR equalizers with different optimality properties. Among these, we formulate the calculation of risk sensitive or minimum entropy FIR filter as the constrained analytic centring problem and mixed H^2/H^∞ problem as another SDP. We provide an example to illustrate the procedures we described.

1. INTRODUCTION

There are various possible reasons for the preference of FIR equalizers over the IIR filters. IIR filters may suffer from the limit cycles caused by the finite precision implementation in real systems and the recursive structure of the IIR filter. Furthermore, the majority of the adaptive equalization approaches makes the FIR assumption and that generally provides with the property that the cost function (either stochastic or deterministic) is a convex function of the FIR equalizer coefficients and has therefore, a single globally optimum solution.

Figure 1 shows the basic structure for the FIR equalization problem. Here $\{b_i\}$ is transmitted sequence, where $b_i \in \mathcal{C}^M$, $H(z)$ is the $N \times M$ transfer function representing the linear distortion effect of the communication channel,$\{v_i\}$ is the noise signal where $v_i \in \mathcal{C}^N$. We also assume that $N \geq M$. We can consider N as the number of antennas and M as the number of cochannel users. Our purpose is to design an FIR equalizer $K(z)$ of order $R-1$ to estimate the delayed version of input sequence $\{b_i\}$.

In this paper, we look at the FIR equalization problem from an H^∞ perspective. First problem that we are going to address is the calculation of the optimal γ value, which is the minimum value of the maximum energy gain from the input disturbances b_i, v_i to the output equalization error sequence e_i, for FIR equalizers. We will use the state space representation in conjunction with the KYP Lemma to pose this problem as a convex Semidefinite Programming (SDP) problem. We will use a similar approach to formulate FIR risk sensitive equalization problem as a constrained analytic centering problem, which is another type of convex problem with LMI constraints. We will later show that the mixed H^2/H^∞ FIR equalizer, that is the H^∞ optimal FIR equalizer with the least H^2 cost, can also be calculated using SDP.

This work was supported in part by DARPA through the Department of Air Force under contract F49620-95-1-0525-P00001 and by the Joint Service Electronics Program at Stanford under contract DAAH04-94-G-0058-P00003.

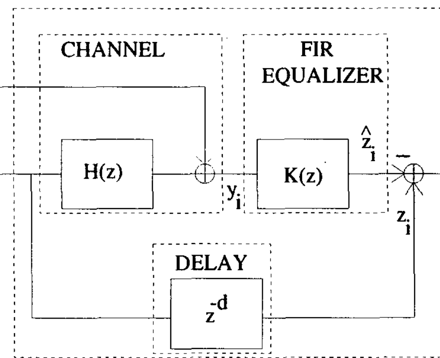

Figure 1: Error Transfer Function

2. STATE SPACE DESCRIPTION FOR THE ERROR TRANSFER FUNCTION

In Section 3, we will use the state space description of the error transfer function T_K, which maps $\begin{bmatrix} b_i & v_i \end{bmatrix}^T$ to e_i, in conjunction with the KYP Lemma to calculate optimal γ value for the H^∞ optimal FIR equalizers and later to calculate equalizer coeffecients themselves.

We will assume that some minimal state space description is given for the communication channel where $\begin{bmatrix} b_i & v_i \end{bmatrix}^T$ is the input and y_i is the output. We can also give a state space description for the delay operator whose dimension depends on delay d.

For the FIR equalizer $K(z) = k_0 + \ldots + k_{R-1}z^{-(R-1)}$ with order $R-1$, we assume the following state space structure

$$\begin{aligned} \xi_{i+1} &= \underbrace{\begin{bmatrix} 0 & 0 \\ I_{(R-2)\cdot N} & 0 \end{bmatrix}}_{F_e} \xi_i + \underbrace{\begin{bmatrix} I \\ 0_{(R-2)\times N} \end{bmatrix}}_{G_e} y_i \\ \hat{z}_i &= \underbrace{\begin{bmatrix} k_1 & \ldots & k_{R-1} \end{bmatrix}}_{H_e} \xi_i + k_0 y_i \end{aligned}$$

Given these, we can obtain the state space structure $\begin{bmatrix} F & G \\ H & D \end{bmatrix}$ for T_K such that

$$T_K(z) = H(zI - F)^{-1}G + D. \tag{1}$$

Note that only H and D are linear functions of equalizer coefficients.

0-7803-6293-4/00/$10.00 ©2000 IEEE.

3. CALCULATION OF γ_{OPT} FOR FIR EQUALIZERS

In this section, we are going to look at the calculation of the optimal γ value using convex optimization techniques.

First, we formulate the FIR equalization problem as follows

$$\inf_{\mathrm{FIR}_{K(z)}} \max_{w} \sigma_{max}(T_K(e^{jw})) = \gamma_{opt,fir}. \quad (2)$$

Therefore, for a given $\gamma \geq \gamma_{opt,fir}$, for $K(z) = k_0 + k_1 z^{-1} + \ldots + k_{R-1} z^{-(R-1)}$ to be an optimal γ-level H^∞ FIR filter, it should satisfy

$$T_K(e^{jw})^* T_K(e^{jw}) \leq \gamma^2 \qquad \forall w \in [0, 2\pi). \quad (3)$$

Using the state space formulation we defined in the previous section, we can give an equivalent condition to the frequency domain condition of Eq. 3. Basic tool we employ for this purpose is Kalman-Yakubovich-Popov (KYP) Lemma which we state here without proof and refer to [2] for the proof:

Lemma 1 KYP Lemma *Consider the observable pair (A, C). Then the following two statements are equivalent*

1. $S_y(z) \geq 0$ *for all* $z = e^{jw} \notin \lambda(F)$ *where* $S_y(z)$ *is*

$$\begin{bmatrix} C(zI-A)^{-1} & I \end{bmatrix} \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (z^{-1}I - A^*)^{-1}C^* \\ I \end{bmatrix} \quad (4)$$

2. *There exists a Hermitian* Z *such that*

$$\begin{bmatrix} Q - Z + AZA^* & S + AZC^* \\ S^* + CZA^* & R + CZC^* \end{bmatrix} \geq 0 \quad (5)$$

In order to use KYP lemma, we need to convert Eq. (3) into the form of Eq. (4). First step in this process is to use Schur complement which is outlined in the following lemma:

Lemma 2 *The following statements are equivalent*

1. *For* $\gamma \geq 0$

$$T_K^*(e^{jw}) T_K(e^{jw}) \leq \gamma^2 I \qquad \forall w \in [0, 2\pi). \quad (6)$$

2.

$$\begin{bmatrix} I & T_K(e^{jw}) \\ T_K^*(e^{jw}) & \gamma^2 I \end{bmatrix} \geq 0 \qquad \forall w \in [0, 2\pi). \quad (7)$$

Proof: The second statement is true if and only if I and its Schur complement $\gamma^2 I - T_K^*(e^{jw})T_K(e^{jw})$ are non-negative, which is the first statement. ■

In order to put matrix in (7) to the form of $S_y(z)$ in 4, we perform the following steps:

$$\begin{bmatrix} I & T_K(e^{jw}) \\ T_K^*(e^{jw}) & \gamma^2 I \end{bmatrix} =$$

$$\begin{bmatrix} I & H(e^{jw}I - F)^{-1}G + D \\ D^* + G^*(e^{-jw}I - F^*)H^* & \gamma^2 I \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 0 \\ G^* \end{bmatrix}}_{C} (e^{-jw}I - \underbrace{F^*}_{A})^{-1} \underbrace{\begin{bmatrix} H^* & 0 \end{bmatrix}}_{S}$$

$$+ \begin{bmatrix} H \\ 0 \end{bmatrix} (e^{jw}I - F)^{-1} \begin{bmatrix} 0 & G \end{bmatrix} + \begin{bmatrix} I & D \\ D^* & \gamma^2 I \end{bmatrix}$$

$$= \begin{bmatrix} C(e^{jw}I - A)^{-1} & I \end{bmatrix} \begin{bmatrix} 0 & S \\ S^* & \underbrace{\begin{bmatrix} I & D \\ D^* & \gamma^2 I \end{bmatrix}}_{R} \end{bmatrix}$$

$$\begin{bmatrix} (e^{-jw}I - A^*)C^* \\ I \end{bmatrix}.$$

We can use KYP Lemma to conclude that if $\gamma \geq \gamma_{opt,fir}$ then there exists a vector $k = \begin{bmatrix} k_0 & \ldots & k_{R-1} \end{bmatrix}$ and a Hermitian matrix Z such that

$$\underbrace{\begin{bmatrix} -Z + AZA^* & S + AZC^* \\ S^* + CZA^* & R + CZC^* \end{bmatrix}}_{\mathcal{F}(k,Z,\gamma)} \geq 0 \quad (8)$$

Note that $\mathcal{F}(k, Z, \gamma)$ is a LMI in k, Z and γ. Therefore, we can define calculation of the optimal value of $\gamma_{opt,fir}$ as a semi-definite programming(SDP) problem:

minimize γ

subject to $\mathcal{F}(k, Z, \gamma) \geq 0$

Solution both yields the optimal value of $\gamma_{opt,fir}$ and a feasible H^∞ optimal FIR filter $k = \begin{bmatrix} k_0 & \ldots & k_{R-1} \end{bmatrix}^T$.

4. RISK SENSITIVE FIR EQUALIZATION

The set of γ-level H^∞ optimal FIR filters, where $\gamma \geq \gamma_{opt,fir}$ is a convex set which can be written as

$$\mathcal{K}_\gamma = \{k : \exists \text{ a hermitian } Z, \mathcal{F}(k, Z, \gamma) \geq 0\}. \quad (9)$$

All these filters in set $\mathcal{K}_\gamma$ have different optimality properties with respect to different criteria. In applications, we desire the FIR equalizer to have some "average" optimality property besides being H^∞ optimal. Our aim in this section is to come up with such an FIR equalizer k_{RS} which is the member of $\mathcal{K}_\gamma$ with the minimum risk sensitive cost. The resulting filter has also property that it is the minimum entropy FIR filter [2]. In the general H^∞ setup, the central solution is the risk sensitive equalizer, however, it is not necessarily FIR.

For the infinite horizon case, we can formulate the FIR risk sensitive equalization problem as

$$\min_{k \in \mathcal{K}_\gamma} - \int_{-\pi}^{\pi} \log(\det(\gamma^2 - T_K(e^{jw})T_K^*(e^{jw})))dw \quad (10)$$

Note that the minimum entropy cost function is clearly a convex function of the equalizer coefficients. Furthermore, since set $\mathcal{K}_\gamma$ is a convex set, the problem in (10) is a convex optimization problem. In the rest of this section, we will try to convert the cost function involving integral into a more compact expression containing the state space variables defined in the previous sections. We first note that

$$\begin{aligned} &\det(\gamma^2 I - T_K(e^{jw})T_K(e^{jw})) \\ = \; &det(\underbrace{\begin{bmatrix} I & T_K^*(e^{jw}) \\ T_K(e^{jw}) & \gamma^2 I \end{bmatrix}}_{S(e^{jw})}). \end{aligned}$$

Since over the set $\mathcal{K}_\gamma$, $S(e^{jw}) \geq 0$, we can define

$$S(e^{jw}) = \Delta_S(e^{jw}) R_e \Delta_S^*(e^{jw}) \qquad w \in [0, 2\pi), \quad (11)$$

where $\Delta_S(z)$ is a monic, causal and causally invertible matrix and $R_e = R + CPC^* \geq 0$. Here P is the solution of Riccati equation

$$-P + APA^* - (APC^* + S)(R + CPC^*)^{-1}(APC^* + S)^* = 0. \quad (12)$$

The observability of (A, C) implies the existence of P. Since Δ_S is analytic for $|z| \geq 1$, it can be shown that

$$\int_{-\pi}^{\pi} \log(\det(\gamma^2 - T_K(e^{jw})T_K^*(e^{jw})))dw = \log(\det(R_e)). \tag{13}$$

Therefore, we can rewrite the optimization problem of (10) as

$$\begin{array}{c} \min_{k\in\mathcal{K}_\gamma, P} - \log(\det(R + CPC^*)) \\ s.t.\text{Eq. (12) holds.} \end{array} \tag{14}$$

Although this formulation looks more desirable than (10), it contains a nonlinear equality constraint. Our aim is to eliminate this nonlinear constraint. For that purpose, we introduce the following convex optimization problem:

$$\begin{array}{c} \min_{k,Z} - \log(\det(R + CZC^*)) \\ \mathcal{F}(k, Z, \gamma) \geq 0. \end{array} \tag{15}$$

The convex optimization problem in (15) is called "constrained analytical centering" problem which is a special case of more general MAXDET problem [5]. It involves nonlinear convex barrier function as the cost function and the convex LMI constraints, and it can be solved very efficiently using interior point methods[5]. If Z_o of the optimal solution (k_o, Z_o) of problem (15) satisfies the Riccati equation

$$Z_o = \underbrace{AZ_oA^* - (AZ_oC^* + S_o)(R_o + CZ_oC^*)^{-1}(CZ_oA^* + S_o^*)}_{Y} \tag{16}$$

,where R_o and S_o are the values of R and S matrices at the optimal point, then it is easy to see that k_o will be the solution of the problem (14) and therefore (10).

In fact, this is generically the case which can be concluded via use of the following theorem[4]:

Theorem 1 Maximal Hermitian Solution *Consider the basic Discrete Algebraic Riccati (DARE):*

$$X = AXA^* + Q - (S + AXC^*)(R + CXC^*)^{-1}(AXC^* + S)^*$$

ad the discrete Riccati inequality

$$AXA^* + Q - (S + AXC^*)^*(R + HXC^*)^{-1}(S + AXC^*)^* \geq X.$$

Let (A, C) be a detectable pair, R be invertible and assume that there is a hermitian solution $\hat{X}$ of (1) for which $R + C\hat{X}C^* > 0$. Then there exists a unique solution $X_+ = X_+^*$ of (1) such that $R + CX_+C^* > 0$ and $X_+ \geq X$ for all hermitian solutions of (1).

The above theorem directly implies that since (A, C) in (17) is a observable and therefore a detectable pair, the maximality property of Riccati equation solution implies the minimality of the $-\log(\det(R + CZC^*))$ and therefore Z_o should satisfy the Riccati equation (17).

We will know take an alternative route to prove this fact via use of KKT optimality conditions for the optimization problem (15). We begin by introducing the corresponding Lagrangian function as:

$$L(Z, k, W) = -log(det(R + CZC^*)) - Tr(W\mathcal{F}(k, Z, \gamma)) \tag{17}$$

where $W \geq 0$ is the dual Lagrange matrix variable. Assuming strict feasibility and therefore the strong duality condition, complementary slackness implies [1] that at the optimal point (k_o, Z_o, W_o),

$$W_o\mathcal{F}(k_o, Z_o, \gamma) = 0$$

$$= \begin{bmatrix} W_{11,o} & W_{12,o} \\ W_{12,o}^* & W_{22,o} \end{bmatrix} \begin{bmatrix} -Z_o + AZ_oA^* & S_o + AZ_oC^* \\ S_o^* + CZ_oA^* & R_o + CZ_oC^* \end{bmatrix}$$

which further implies,

$$\begin{aligned} W_{12,o} &= -W_{11,o}(S_o + AZ_oC^*)(R_o + CZ_oC^*)^{-1} && (18) \\ W_{22,o} &= (R_o + CZ_oC^*)^{-1}(S_o + AZ_oC^*)^*W_{11,o} \\ &\quad (S_o + AZ_oC^*)(R_o + CZ_oC^*)^{-1} && (19) \\ 0 &= W_{11,o}(-Z_o + Y). && (20) \end{aligned}$$

Note that Eq.(20) implies that if $W_{11,o} > 0$, i.e. strictly positive, then Eq.(17) should hold. In order to check this condition, we use the first order optimality condition of the Lagrangian function.

If we differentiate the Lagrangian function with respect to matrix Z, we obtain

$$\begin{aligned} \nabla_Z L(k, Z, W) &= -C^*(R + CZC^*)^{-1}C + W_{11} \\ &\quad -A^*W_{11}A - A^*W_{12}C - C^*W_{12}^*A - C^*W_{22}C \end{aligned}$$

At the optimal point, using Eqs. (18) and (19)

$$\begin{aligned} \nabla_Z L(k, Z, W)|_{k_o,Z_o,W_o} &= -C^*(R_o + CZ_oC^*)^{-1}C \\ &\quad +W_{11,o} - (A + C^*M_o^*)W_{11,o}(A + MC) \end{aligned}$$

where $M_o = -(S_o + AZ_oC^*)(R_o + CZ_oC^*)^{-1}$. First order optimality condition implies

$$\begin{aligned} \nabla_Z L(k, Z, W)|_{k_o,Z_o,W_o} &= 0 && (21) \\ W_{11,o} - (A + C^*M_o^*)W_{11,o}(A + M_oC) &= N_oN_o^* && (22) \end{aligned}$$

where $N_o = B^*(R_o + CZ_oC^*)^{-1/2}$. Since Eq. (22) is a Lyapunov equation, $W_{11} > 0$ if and only if pair $(A + M_oC, N_o)$ is observable. As a result, since observability of (A, C) pair implies observablity of $(A + M_oC, N_o)$, $W_{11} > 0$, and therefore, Eq. (17) holds.

To conclude, in this section we showed that the risk sensitive or minimum entropy FIR equalization problem can be posed as a finite dimensional convex optimization problem (constrained analytic centering problem) and therefore can be efficiently solved using interior point algorithms.

5. MIXED H^2/H^∞ FIR EQUALIZATION

Using convex optimization techniques, we can also find solution to mixed H^2/H^∞ FIR equalization problem. This problem refers to finding the H^∞ optimal FIR equalizer which has the least H^2 norm and it can be posed as an SDP.

The following lemma[3], which provides the calculation of H^2 norm as a function of state space parameters, plays a central role in the SDP formulation of the mixed problem:

Lemma 3 *(**H^∞ norm bound**) Given any transfer function $H(z) = C(zI - A)^{-1}B + D$ (not necessarily minimal), we have:*

$$||H(z)||_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} Tr\left(H(e^{jw})H^*(e^{jw})\right)dw < \beta^2$$

where A is asymptotically stable, if and only if the following LMI in X and S is feasible:

$$\begin{bmatrix} A^*XA - X & A^*XB \\ B^*XA & B^*XB - I \end{bmatrix} < 0$$

$$\begin{bmatrix} X & 0 & C^* \\ 0 & I & D^* \\ C & D & S \end{bmatrix} > 0$$

$$Tr(S) - \beta^2 < 0$$

$$X > 0.$$

Therefore minimizing β^2 under above constraints together with $F(k, Z, \gamma) \geq 0$ would yield the mixed H^2/H^∞ solution. Similar to risk sensitive FIR filter, mixed H^2/H^∞ solution has a desired average property which is minimization of mean square error under some statistical assumptions about the disturbances.

In the next section, we will provide comparison of FIR filters that are optimal with respect to different criteria for an example setting.

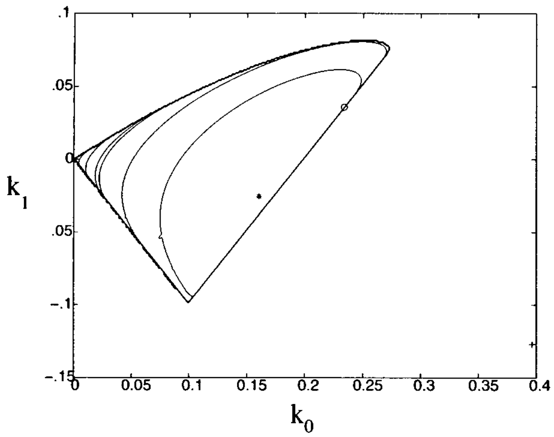

Figure 2: k_0 vs. k_1. Two tap equalizers for $H(z) = 1 + 0.9z^{-1}$ and $d = 0$: '*'- Risk Sensitive Solution, '+'- H^2 Solution, and 'o'- Mixed H^2/H^∞ Solution

6. CONCLUSION AND AN EXAMPLE

In this paper, using state space based approach, we formulated calculation of FIR H^∞ equalizers with good "average" properties as convex optimization problems. In particular, we showed risk sensitive (or minimum entropy) equalization can be formulated as constrained analytic centring problem and mixed H^2/H^∞ equalization as SDP, which can be solved using efficient convex optimization algorithms.

In order to illustrate the methods we presented, we consider the channel $H(z) = 1 + 0.9z^{-1}$ and the delay $d = 0$ as a simple example. In order to obtain a geometrical picture for the set of equalizers, we consider the equalizers of length 2. In Figure 2, the bounded convex region, the spectrahedron, represents the set $\mathcal{K}_\gamma$, γ-level H^∞ optimal equalizers, for this setting. The contours for the risk sensitive cost function are also drawn inside the spectrahedron. The point represented by '*' is the risk sensitive equalizer which is obtained by the constrained analytic centering method described previously. The point marked by '+' is the H^2 solution which is apparently not H^∞ optimal since it lies outside of the set of H^∞-optimal equalizers. The point marked with 'o' is the mixed H^2/H^∞ solution. Since the H^2-optimal equalizer is not inside the spectrahedron of H^∞ solutions, the mixed H^2/H^∞ solution lies at the boundary where the H^2 cost function is minimum.

In Figure 3, the error spectra for the equalizers of the above setting are shown. It is clear from this figure that, since both the risk sensitive and mixed H^2/H^∞ solutions are H^∞ optimal, their error spectra have the same maximum value which is smaller than the the maximum value of the H^2 equalizer's error spectrum. The total area under the error spectrum is clearly minimized by the H^2 solution as expected. Besides, the mixed H^2/H^∞ solution has smaller are under the error spectrum than the risk sensitive case, however, error spectrum for the risk sensitive equalizer is smaller than the mixed H^2/H^∞ solution at most of the frequencies, especially around the frequencies where the error spectra for both have high values.

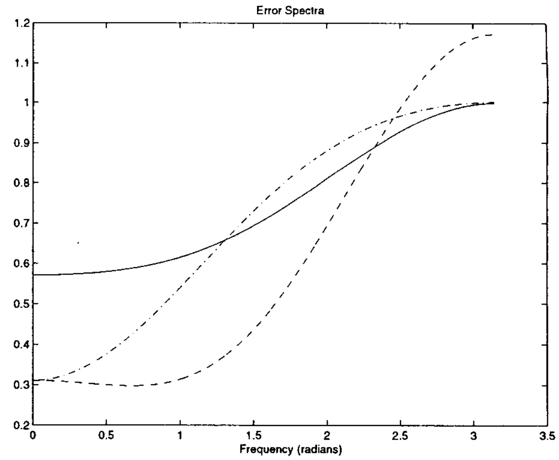

Figure 3: The Error Spectra for two tap equalizers for $H(z) = 1{+}0.9z^{-1}$ and $d = 0$: 'solid'- Risk Sensitive Solution, 'dashed'- H^2 Solution, and '-.'- Mixed H^2/H^∞ Solution

7. REFERENCES

[1] F. Alizadeh, J. Haeberly, and Michael Overton. Complementarity and nondegeneracy in semidefinite programming. *Submitted to Math. Programming*, 1995.

[2] B. Hassibi, A.H. Sayed, and T. Kailath. *Indefinite Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories*. SIAM Studies in Applied Mathematics, New York, 1998.

[3] Haitham Hindi, Babak Hassibi, and Stephen Boyd. Multiobjective H^2/H^∞-Optimal control via finite dimensional Q-Parametrization and linear matrix inequalities. *Proceedings of American Control Conference*, 5:3244–9, 1998.

[4] Peter Lancaster and Leiba Rodman. *Algebraic Riccati Equations*. Oxford Science Publications, 1995.

[5] S. Wu. Max-det programming with applications in magnitude filter design. *Ph.D. Thesis, Stanford University*, 1997.